AD A049590

AD No.
DDC FILE COPY

COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

DDC
RECEIVED
FEB 8 1978
A

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

(1)

(6)

ALGORITHMS AND HARDWARE TECHNOLOGY FOR IMAGE RECOGNITION

(9) Third Semi-Annual Report. no. 3,
1 May ~~1977~~ - 31 October 1977,

(15) Contract DAAG53-76C-0138,
(DARPA Order-3206)

(11) Nov 77

Computer Science Center
University of Maryland
College Park, MD 20742

(12) 27p.

ABSTRACT

The research activities summarized in this report relate to the development of algorithms for automatic target cueing in FLIR imagery. Topics covered include preprocessing and feature extraction (noise cleaning, edge detection); using image models to predict the outputs of thresholding and edge detection operations; segmentation and postprocessing (variable thresholding, adaptive quantization, thinning, recursive segmentation); and classification. Westinghouse, under a subcontract, has studied implementation of selected algorithms in CCD circuitry; this work is described in separate reports.

DDC
RECEIVED
FEB 8 1978
A

The support of the U. S. Army Night Vision Laboratory under Contract DAAG53-76C-0138 (DARPA Order 3206) is gratefully acknowledged, as is the help of Mrs. Shelly Rowe in preparing this report.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

403 018

TABLE OF CONTENTS

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

1. Preprocessing and feature extraction

1.1 Precleaning: min/max processing

The operations of local maximum and minimum commute with any monotone transformation of the gray level. In image processing, many gray level mappings are monotone functions. In particular, thresholding commutes with max and min. Thus the sequence of shrink and expand noise cleaning operations normally applied to the (binary) results of thresholding may be viewed as a sequence of min and max operations applied to a gray level image. By commutativity, the sequence may be applied prior to thresholding. The "precleaned" image which results is smoother than the original image. Its histogram is somewhat improved as well (e.g., better peak separation), facilitating segmentation. The utility of precleaning is that no early commitment to a threshold need be made. In fact, if several thresholds are used (as might occur in the Superslice algorithm), then separate noise cleaning passes are unnecessary. This work is discussed more completely in a separate report [1].

1.2 Predicting results of thresholding

Thresholding images is a common process and much work has been directed towards selecting the correct value at which to threshold and estimating the expected error. Normally, one thresholds only images which contain some signal. Thresholding pure noise is to be avoided when possible. Since there may be occasions when thresholding noise is unavoidable (e.g., a poor threshold was chosen), it is important to predict the expected results. Thus a bad threshold will often cause an object region to break up into smaller noise regions. The expected number of such noise regions is useful in planning for data structure storage allocation. There are no current methods for predicting the number of connected components of thresholded spatially correlated signal (or noise). However, recent work [2] makes it possible to estimate the moments of regions, the density of border points, and lower bounds on the number of connected components in thresholded noise images. The input grayscale image is modeled as a two-dimensional random process characterized by its mean and power spectrum. Tests with synthetic and actual data compare the predicted and measured responses. The comparison shows reasonable agreement between the two, although the predictions are worst for thresholds at or near the mode of the noise distribution.

1.3 Predicting edge detector response

Statistical response prediction for edge operators can be used to determine the nature of further processing of the response. If edge detector output is to be thinned or thresholded, the false alarm and false dismissal rates depend on the statistics of the operator responses. A technical report [3] has been completed which discusses the statistical properties of the outputs of some edge detectors operating on a general class of images.

The image model considered is the same as in the previous section. Thus we wish to consider the response of edge detectors to background noise. The edge detectors analyzed are the Laplacian and its absolute value, and the absolute difference of adjacent averages over 2x2 and 4x4 neighborhoods. The response features which were measured are the mean edge response at each point, the variance, the auto-covariance, and the cross-covariance of gray level and edge value at a number of displacements. In addition, the density of the local maxima of edge values is computed. A set of synthetic background images showed good conformity to the predicted features.

1.4 Optimal edge detection

Many optimality criteria have been proposed for edge detection. Among the most well known is that devised by Hueckel [4]. It involves fitting an ideal parameterized step edge to the image data so as to minimize the mean squared error. In a recently issued report [5], a new optimal detector simplifies several assumptions associated with the Hueckel detector and thereby solves an easier optimization problem. Specifically, by assuming that the local mean is zero and the local variance is unity, two Hueckel parameters can be eliminated. Further simplifications follow if the operator can be applied at each point with the edge assumed to pass through the center (or not to exist at all). The resulting formulation can be tuned to favor edges with known a priori probabilities. The computational effort involved in applying the operator may be reduced by solving the associated cubic equation using a simple iterative approximation, such as Newton's formula. Further testing on actual data is necessary in order to establish the utility of this approach.

1.5 Evaluation of edge operators

Relaxation and other intermediate-level edge extraction processes are able to utilize more information than simply the edge response at each point. With this in mind, it seems appropriate to evaluate a number of edge operators for their adequacy in obtaining the magnitude, direction, and reliability of the edge response at some set of image points, for both ideal and distorted images.

Among the operators particularly appropriate for this use, one introduced by Hueckel in 1973 was found, in the course of the investigation, to incorporate a theoretical flaw leading to eccentric behavior in textured images. An operator which is conceptually similar but apparently more dependable has now been defined, and is being tested along with several others including Hueckel's operator, a modified Hueckel form suggested by Meró and Vássy, a local operator recently described by Hummel, and the Sobel operator. Comparative testing is currently in progress.

1.6 Conformity

An algorithm has been developed, in conjunction with the Superslice algorithm, which measures the extent to which a subset of the border of a region surrounds the region. Let B be the set of border points of a region R. Let $E \subset B$ be a subset. The natural ordering of B induces an ordering on E. Thus E determines a unique polygon Q (just as B determines R). If Q is identical to R then E is a "best" subborder of R. Naturally, E may be non-existent or non-unique. If Q is not identical to R then the goodness of fit of Q to R (as measured by the area of $(Q \cap \bar{R}) \cup (\bar{Q} \cap R)$) defines a measure of "conformity" of E to B. This is a more sophisticated use of border/edge cooccurrence than simply counting the percentage of matched border points of a region. The conformity measure is computed during the same pass as the other measures of object region definition in Superslice. Scatterplots of target and noise regions for the conformity feature plotted against contrast were compared with similar plots for the edge match feature. The target clusters were tighter in the conformity/contrast plots, suggesting that conformity is a better discriminator than edge match.

2. Segmentation and Postprocessing

2.1 Variable thresholding

If the background gray level in an image varies gradually across the image, it may not be possible to extract objects by applying a single threshold to the entire image, even if the objects are all of the same type. In such situations, a method of variable thresholding proposed some years ago by Chow and Kaneko [6], in connection with segmentation of chest x-rays, can be used. The image is divided into windows, a histogram is computed for each window, and a bimodality test is applied to each histogram. (Some windows will have bimodal histograms, since they contain objects or parts of objects surrounded by background, while others will have unimodal histograms because they contain object points only or background points only.) For each bimodal histogram, an appropriate threshold is selected. This yields a set of window thresholds that are applicable to various parts of the image. The remaining parts of the image can be thresholded by interpolating (or extrapolating) on this set of window thresholds.

This variable thresholding scheme is in process of being implemented. It is planned to test it in conjunction with the Superslice algorithm, as an alternative to applying a set of thresholds to the whole image.

2.2 Histogram sharpening and adaptive quantization

If an image can be modelled as the spatial concatenation of homogeneous "patches" of different gray levels, the gray level histogram will appear as a sequence of peaks separated by valleys. In practice, the peaks tend to appear somewhat smeared. One can sharpen the histogram by identifying significant peaks and agglomerating those points whose gray levels are on the slopes of the same peak. This requantization process has a number of beneficial results. First, the resulting images look remarkably like the original images. Second, they typically have about 1/10 the number of gray levels. This is a significant compression factor considering that no spatial information is utilized. Third, the isolated peaks in the transformed histograms induce a convenient, primitive segmentation of the original image. Further details on this work can be found in [7].

2.3 Clustering contour points

Slice-level selection is a topic of continuing interest. Our approach is to produce clusters of points corresponding to region borders and to associate the average gray level of each cluster with a threshold for the corresponding region. Region borders usually correspond to points of locally maximum edge response. Previous work suggests the use of edge detectors which select at each point the maximum difference of averages of adjacent neighborhoods over several directions. By suppressing non-maximum responses normal to the selected direction (i.e., across the edge), thin contours result which appear to surround object regions.

This process produces as a by-product points with very low edge value, including values which truncate to zero. Such points correspond to the interiors of homogeneous regions. It is useful to include such points in our analysis, even though they constitute a population with fundamentally different properties from contour points.

Once the thinning has been accomplished, the resulting points can be accumulated in a two-dimensional histogram. Each point contributes an edge value and a gray level value. In practice, it was found that the gray level of the point does not serve as a good coordinate. Specifically, for a step edge, the gray level of an edge point lies in one population or the other but not at any intermediate gray value.

Thus the histogram consists of a pair of spikes corresponding to the gray levels on either side of the step edge. A single cluster results if the average gray level of the two neighborhoods which contributed the maximal edge response at the point is plotted on the gray scale.

The relation of edge clusters to interior clusters gives rise to a segmentation strategy based on partitioning the 2-D histogram into disjoint regions representing object classes. However, its dependence on the accuracy of cluster extraction limits its utility for the moment.

2.4 A Recursive segmentation algorithm

The Superslice algorithm [8] has been combined with Ohlander's approach to segmentation [9] to provide a more powerful method of extracting regions from an image. Currently operating in an interactive mode, the program ultimately would run as follows:

RECURSIVE PROGRAM SEGMENT(PIC)

CHOOSE THE BEST SLICE RANGE (LO,HI) FOR PIC (CF. SECTION 2.4)

IF NO SLICE RANGE EXISTS (I.E., PIC IS UNIMODAL)

   THEN OUTPUT PIC AS AN ATOMIC REGION AND RETURN

   ELSE SLICE PIC AT (LO,HI) TO PRODUCE SPIC END

DO SUPERSLICE ON SPIC AND EXTRACT REGIONS $R_1,\ldots,R_K$ END

IF K=0 THEN OUTPUT SPIC AS A (TEXTURED) REGION AND RETURN END

FOR I=1 TO K DO BORDER CLEANING ON R_I; CALL SEGMENT (R_I) END

LET RPIC = PIC - $\bigcup_{I=1}^{K} R_I$

CALL SEGMENT (RPIC)

RETURN

END

2.5 Fuzzy thinning

Objects which are everywhere elongated are often thinned down to a "medial line" for the purpose of extracting thickness-invariant topological features of the objects. The basic strategy for thinning is to iteratively delete endpoints or border points of an object which do not locally disconnect it. For binary images, various parallel algorithms exist. The recent extension of the topological concept of connectedness to fuzzy subsets allows us to generalize thinning to gray level images. Assume we have thin dark objects on a light background. We define gray level thinning to be the successive replacement of points by the minimum gray level of their neighbors if those changes do not affect the local fuzzy connectedness for any pair of neighbors. The result of applying such an algorithm is a set of high gray level "curves" lying on the ridges and peaks of high gray level in the original picture. If the original picture is noisy there will be many local peaks; so while thinning is defined for unsegmented pictures, a local threshold is necessary to overlook these small noise peaks. Unlike binary thinning, however, we no longer need to distinguish between border and interior points since thinning a homogeneous region will not significantly change the gray level of any point; only a slight smoothing results. The results of experiments with this technique are described in a technical report [10].

3. Classification

3.1 Object classification - preliminary results

The basic procedure being used to classify extracted regions has been described in earlier reports. A discussion of the relative utility of the features used in classification, as well as an overall review of the classification results, is in preparation. The general result, however, is that target/noise discrimination is quite good. A recent test, designed to be directly comparable with a Honeywell test on the same images, is described below, including summary confusion matrices for two classification experiments and corresponding data for similar Honeywell experiments.

The first test used the entire set of (large object) images to train a (4-class) maximum-likelihood classifier, then used that classifier on the same images. Six features were used (four shape features and two edge features), compared to six shape discriminants used in the matching Honeywell test. The confusion matrices are:

	"classified"	
"true"	targets	non-targets
targets	107	4
non-targets	7	25

Maryland study

self-classification:all targets

	"classified"	
"true"	targets	non-targets
targets	59	10
non-targets	18	113

Honeywell study

self-classification:all extracted targets, with selected non-targets included.

The differences apparent in total number of targets reflect the far lower success rate of target extraction in the Honeywell study, and a rigorous pre-filtering of noise regions before classification in the Maryland scheme, which rejects many non-targets before the classification stage is reached. Because of this, it is the total number of misclassifications, rather than percentage of misclassification, which is comparable.

As the objects were extracted on rather different bases, however, it may be helpful to express the latter values as rates per frame. Twenty-five windows (each 64x64 pixels, sampled from a 128x128 median filtered window) in the Maryland study cover approximately the area of one 500x800 pixel frame. If the number and characteristics of the non-target objects extracted were given throughout the frame by the rates for the target windows, the per frame false-alarm rate, due to 'large' noise objects, would be:

$$\frac{7 \text{ false alarms}}{150 \text{ windows}} \times \frac{25 \text{ windows}}{\text{frame}} \sim 1.2 \text{ false alarms/frame}$$

The corresponding number from the Honeywell study is (assuming 50% of all non-targets are 'large')

$$\frac{18 \text{ false alarms}}{131 \text{ non-targets}} \times \frac{600 \text{ non-targets}}{110 \text{ frames}} \sim 0.75 \text{ false alarms/frame}$$

[The corresponding rate for small targets in the Honeywell

study is ∿ 1.4 false alarms/frame.]

In the second test, the data were split into two approximately equal populations. A classifier was trained on one set of objects, then used to classify both sets. The confusion matrices obtained are:

	"classified"	
"true"	targets	non-targets
targets	53	1
non-targets	3	13

Training set

	"classified"	
"true"	targets	non-targets
targets	53	4
non-targets	6	10

Test set

Maryland study

	"classified"	
"true"	targets	non-targets
targets	28	7
non-targets	6	159

Training set

	"classified"	
"true"	targets	non-targets
targets	16	20
non-targets	6	160

Test set

Honeywell study

Notice that the Maryland classification results for training and test samples indicate more reliable estimates of the error rates, as well as a much lower false dismissal rate even after extraction (4.5% (Md.) versus 38% (Honeywell)). Calculating per frame rates as above, we would get for Maryland (for large non-targets):

$$\frac{9}{150} \times 25 \sim 1.5 \text{ false alarms/frame,}$$

and for Honeywell:

$$\frac{12}{331} \times \frac{1200}{110} \sim 0.4 \text{ false alarms/frame}$$

[Again, Honeywell's corresponding small-target rate was $\sim$ 0.6 false alarms/frame, while Maryland's is not yet available.]

3.2 Class structure improvement

While the classification scheme presently being used seems satisfactory, it does require statistical characterization of "non-targets" as a single class, or set of classes. More logical, and in some sense more robust, would be a classifier which characterized the classes of real interest (here: tank, APC, and truck) and simply rejected any object which was not sufficiently like any of the target classes. Furthermore, the notion of allowing the classifier an output of "target;class unknown" also seems appropriate. The resultant scheme is basically a normal statistical classifier of the type already described, but with a final test to determine the confidence in the classification, (e.g., by the introduction of two "reject" zones: "non-target" and "unknown target"). Design and testing of this classifier is in progress.

3.3 Low-level context for classification validation

Our approach to the target cueing problem has been to extract and classify object regions independently of one another. That is, segmentation is based on the assumption that object regions are individually thresholdable, though not necessarily all by the same threshold. Classification is based on information derived from measurements on the individual components, but does not take into account the intra- and inter-frame context of a region with all others. The constraints operating between regions define a kind of second order relation on parts of a frame which could be used to verify or revise the decisions of the classifier.

The Gestalt rules of grouping are of interest in this respect since they refer to factors that cause some parts to be seen as belonging more closely together than others. These rules are applications of the basic "principle of similarity" which asserts that region association is partly defined by region resemblance.

There are several types of similarity which are of interest for FLIR imagery. First, similarity of appearance groups together regions whose attributes of size, shape, brightness, etc. are similar. That is, regions which are close in feature space, regardless of their locations with respect to the discriminant surfaces, look alike. Therefore, a classifier which identifies such regions as separate target types should consider alternative information sources to

resolve these possible inconsistencies.

Similarity of location or proximity produces clusters of regions which are physically near one another. If the segmentation procedure has fragmented a target, e.g., by separating the barrel from the body of a tank, then it may be of value for the classifier to know about all the components, irrespective of their characteristics, which are proximate. In this way small objects, which are not significant on their own, may prove to be important when viewed as a part of a larger object.

Similarity by spatial arrangement organizes parts according to their geometric configurations. General knowledge about patterns of target behavior, i.e., tanks tend to travel in collinear groups of three to five, could be used to verify the results of the classifier on a given object region when the types and locations of other targets in the frame suggest it.

Temporal similarity groups inter-frame object regions utilizing the heuristic that region descriptions of the same object tend to cluster more closely than do descriptions of different objects. Furthermore, we can use these multiple descriptions of the same object to derive a best exemplar for classification. A method of tracking objects using dynamic programming has been described in previous progress reports.

The principles listed above constitute a low-level context for object regions, and introduce potentially powerful semantics into the classification problem.

4. Hardware

4.1 Sorter chip fabrication

During this period, Westinghouse fabricated and successfully demonstrated a chip embodying a sorter function in CCD architecture. This sorter function is a fundamental primitive which is used in the median filter, the histogrammer, and in ordering region descriptions by significance. The chip operates within the time, space and cost constraints established early in the project. Further details can be found in Westinghouse's Fifth and Sixth Quarterly Reports.

4.2 Connected component algorithm

Westinghouse has completed the design of the connected component coloring algorithm, a major part of the Superslice segmentation scheme. The design of an implementation of this algorithm in CCD architecture will take place in the final quarter.

Bibliography

1. Nakagawa, Y., and A. Rosenfeld, A note on the use of local min and max operations in digital picture processing, University of Maryland Computer Science Center Technical Report 590, October 1977.

2. Panda, D., Statistical properties of thresholded images, University of Maryland Computer Science Center Technical Report 559, August 1977.

3. Panda, D., Statistical analysis of some edge operators, University of Maryland Computer Science Center Technical Report 558, August 1977.

4. Hueckel, M. H., A local visual operator which recognizes edges and lines, JACM, V20, 1973, pp. 634-647.

5. Hummel, R., Edge detection using basis functions, University of Maryland Computer Science Center Technical Report 569, August 1977.

6. Chow, C. K., and T. Kaneko, Automatic boundary detection of the left ventricle from cineangiograms, Computers in Biomedical Research, V5, 1972, pp. 388-410.

7. Peleg, S., Iterative histogram modification, 2, University of Maryland Computer Science Center Technical Report 606, November 1977.

8. Milgram, D., Region extraction using convergent evidence, Proceedings: Image Understanding Workshop, April 1977, Minneapolis, MN, pp. 58-64.

9. Ohlander, R. B., Analysis of natural scenes, Technical Report, Computer Science Department, Carnegie-Mellon University, 1975.

10. Dyer, C., and A. Rosenfeld, Thinning algorithms for gray-scale pictures, University of Maryland Computer Science Center Technical Report 610, November 1977.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ALGORITHMS AND HARDWARE TECHNOLOGY FOR IMAGE RECOGNITION Third Semi-Annual Report 1 May 1977 - 31 October 1977		5. TYPE OF REPORT & PERIOD COVERED Semi-Annual 5/1/77 - 10/31/77
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s) DAAG53-76C-0138 (DARPA Order 3206)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Center University of Maryland College Park, MD 20742		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE November, 1977
		13. NUMBER OF PAGES 24
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) U.S. Army Night Vision Laboratory Fort Belvoir, VA 22060		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image understanding
Image processing
Pattern recognition
Target detection
FLIR imagery

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The research activities summarized in this report relate to the development of algorithms for automatic target cueing in FLIR imagery. Topics covered include preprocessing and feature extraction (noise cleaning, edge detection); using image models to predict the outputs of thresholding and edge detection operations; segmentation and postprocessing (variable thresholding, adaptive quantization, thinning, recursive segmentation); and classification. Westinghouse, under a subcontract, has studied implementation of

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. selected algorithms in CCD circuitry; this work is described in separate reports.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)